


(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
n: Wed Jul 22 15:27:41 1998; MasPar time 12.38 Seconds
767.369 Million cell updates/sec
Tabular output not generated.

Title: >US-08-915-659A-7
Description: (1-260) from US08915659A.pep
Perfect Score: 1943
Sequence: 1 MGRPRPRAAKTWMLLLGG.....VYTNICRYLDWIKKIIGSKG 260

Scoring table: PAM 150
Gap 11

Searched: 120441 seqs, 36531193 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: p156
1:p1r1 2:p1r2 3:p1r3 4:p1r4 5:nrl3d

Statistics: Mean 42.128; Variance 70.703; scale 0.596

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Query	Score	Match	Length	DB	ID	Description	Pred. No.
1	1488	76.6	260	2	I56559	neuropsin - mouse	0.00e+00
2	833	42.9	253	2	A53968	serine proteinase SCC	3.84e-170
3	808	41.6	247	1	TRDG	trypsin (EC 3.4.21.4)	5.44e-164
4	804	41.4	238	2	S31779	trypsin (EC 3.4.21.4)	5.23e-163
5	795	40.9	246	1	TRDGC	trypsin (EC 3.4.21.4)	8.52e-161
6	785	40.4	229	1	TRBOPR	trypsin (EC 3.4.21.4)	2.43e-158
7	783	40.3	247	2	A27547	trypsin (EC 3.4.21.4)	7.53e-158
8	781	40.2	246	1	TRRT2	trypsin (EC 3.4.21.4)	2.33e-157
9	779	40.1	225	5	ITGS2	trypsin (EC 3.4.21.4)	7.22e-157
10	768	39.5	223	5	2TLDE	trypsin (EC 3.4.21.4)	3.60e-154
11	768	39.5	246	2	B25528	trypsin (EC 3.4.21.4)	3.60e-154
12	767	39.5	246	1	TRRT1	trypsin (EC 3.4.21.4)	6.33e-154
13	766	39.4	223	5	ITPS	trypsin complexed wit	1.11e-153
14	766	39.4	223	5	ITYN	Beta trypsin complexe	1.11e-153
15	766	39.4	223	5	IURTA	trypsin (EC 3.4.21.4)	1.11e-153
16	766	39.4	223	5	IURSA	trypsin (EC 3.4.21.4)	1.11e-153
17	766	39.4	223	5	ITGB	trypsin (EC 3.4.21.4)	1.11e-153
18	766	39.4	223	5	ITGT	trypsin (EC 3.4.21.4)	1.11e-153
19	766	39.4	223	5	4PTPIZ	trypsin (EC 3.4.21.4)	1.11e-153
20	766	39.4	223	5	2PTN	trypsin (EC 3.4.21.4)	1.11e-153
21	766	39.4	223	5	ITPP	trypsin (EC 3.4.21.4)	1.11e-153
22	766	39.4	223	5	ITBO	trypsin (EC 3.4.21.4)	1.11e-153
23	766	39.4	223	5	IBTY	beta-trypsin (EC 3.4.21.4)	1.11e-153

ALIGNMENTS

```
RESULT 1
ENTRY 1
TITLE I56559 #type complete
ORGANISM neuropsin - mouse
DATE 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 08-Sep-1997
ACCESSIONS I56559
REFERENCE I56559
#authors Ched Z.L.; Yoshida, S.; Kato, K.; Momota, Y.; Suzuki, J.; Tanaka, T.; Ito, J.; Nishino, H.; Aimoto, S.; Kiyama, H.; J. Neurosci. (1995) 15:5088-5097
#journal Expression and activity-dependent changes of a novel limbic-serine protease gene in the hippocampus.
#title
#cross-references MUID:95348817
#accession I56559
#status Preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-260 #label RES
#cross-references GB:D30785; NID:g1020090; PID:g1020091
CLASSIFICATION #superfamily trypsin; trypsin homology
FEATURE
33-252 #domain trypsin homology #label TRY
SUMMARY #length 260 #molecular-weight 28523 #checksum 6448
```

Query Match	76.6%	Score 1488;	DB 2;	Length 260;
Best Local Similarity	72.2%	Pred. No. 0.00e+00;		
Matches	187;	Conservative 38;	Mismatches 34;	Indels 0; Gaps 0;
Db	1	MGRPPCAIQPWTLLLFMGAWAGLTRAQGSKTLREGICIPHSPQWQAALFOGERLICGG	60	
QY	1	MGRPRPRAAKTWMLLLGGAWAGHSAQEDKVLGHECQPHSQWQAALFOGQQLCGG	60	
Db	61	VLYGDRWLTAACHCKQKYSVRLGDSLSRQDPEBIEIQAQSIQHPCTYNNPDEHSD	120	
QY	61	VLYGGNWLTAACHCKPKYTVRLGDSLSQNKDGPQEIPIVYQSIHPCTYNNPDEHSD	120	
Db	121	IMLIRLQNSANLGDKVKPVOLANLCPKVGOKCTIISGWTGVTSPQENFPNTLCAEYKIYS	180	
QY	121	IMLILQRLQNSLGSKVKPISLADHCTQPGQKCTVSGWGVITSPRENFPTLCAEYKIFP	180	
Db	181	QNKCERAYPKITEGMVCAAGSSNGADTCQDSSGGLVCDGMLQGITSWGSDPGCKPEKPG	240	
QY	181	QKCEDAYPQITDGMVCAAGSSNGADTCQDSSGGLVCDGMLQGITSWGSDPGCKPEKPG	240	
Db	241	VYTKICRYTWIKTKWMDNR	259	
QY	241	VYTNICRYLDWIKKIIGSKG	259	


```

QY 190 QG1TDMGVCASSKA-DTCOGDSGLPVCDAALQGTSGSDPCGRSDKPGVYTNICRY 248
Db 236 VSWIQOTIAAN 246
: |||
QY 249 LDWIKKLGSK 259

RESULT 6
ENTRY TRBOTR #type complete
TITLE trypsin (EC 3.4.21.4) precursor - bovine
CONTAINS trypsinogen
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE 24-Apr-1984 #sequence_revision 28-Feb-1986 #text_change
18-Jul-1997
ACCESSIONS A90164; A00946; S08774
REFERENCE A90164
#authors Mikes, O.; Holeysovsky, V.; Tomasek, V.; Sorm, F.
#journal Biochem. Biophys. Res. Commun. (1966) 24:346-352
#title Covalent structure of bovine trypsinogen. The position of the
remaining amides.
#cross-references MWID:67168848
#accession A90164
#molecule_type protein
#residues 1-57,'Q',59-67,'Q',69-150,'N',152-176,'N',178-229
##label MIK

REFERENCE A93755
#authors Hartley, B.S.
#journal Philos. Trans. R. Soc. Lond. (1970) B257:77-87
#contents annotation; revisions
REFERENCE A00950
#authors Titani, K.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.
#journal Biochemistry (1975) 14:1358-1366
#title Amino acid sequence of dogfish trypsin.
#cross-references MWID:7514645
#contents annotation; revisions
#note the sequence agrees with that shown
REFERENCE A92954
#authors Bode, W.; Schwager, P.
#journal J. Mol. Biol. (1975) 98:693-717
#title The refined crystal structure of bovine beta-trypsin at 1.8
angstrom resolution.
#cross-references MWID:76072097
#contents annotation; X-ray crystallography; binding sites for calcium,
substrate, and inhibitors
COMMENT Trypsinogen is synthesized in the acinar cells of the pancreas.
COMMENT Autocatalytic cleavage after Lys-6 leads to beta-trypsin by
releasing a terminal hexapeptide. Subsequent cleavage after
Lys-131 leads to alpha-trypsin. Further cleavage after Lys-176
yields pseudotrypsin. A cleavage may also occur after Arg-105.
CLASSIFICATION #superfamily trypsin; trypsin homology
KEYWORDS hydrolase; pancreas; protein digestion; serine proteinase;
zymogen

FEATURE
1-229 #product trypsinogen #status experimental #label 2YM\
1-6 #domain activation peptide #status experimental #label
APT\
7-222 #domain trypsin homology #label TRV\
7-131,132-229 #product alpha-trypsin #status experimental #label MPT\
6-7 #cleavage_site Lys-Ile (enteropeptidase) #status
experimental\
13-143,31-47, #disulfide_bonds #status experimental\
115-216,122-189, #active_site His, Asp, Ser #status experimental\
154-168,179-203 #binding_site calcium (Glu, Asn, Val, Glu) #status
46,90,183 experimental\
58,60,63,68 #cleavage_site Lys-Ser (autolytic) #status experimental
131-132 #length 229 #molecular_weight 23993 #checksum 2248
SUMMARY

Query Match 40.4%; Score 785; DB 1; Length 229;
Best Local Similarity 46.6%; Pred. No. 2.43e-158;
Matches 108; Conservative 50; Mismatches 58; Indels 6; Gaps 6;

```

```

Db 3 DDDKIVGGYTCGANTVPYQVSLNSGYHF-CGSLINQWVYVAAHCKYSGIQVRLGEDNI 61
Qy 29 QEDVLGGHECOPHSQWQAALFOGQQLLGGVVGWVLTAAHCKPKYTVRLGDHSL 88
Db 62 NVVSGNQFISAKSIYHPSYNSNTL-N-N-DIMLIKLSAASLSNRVASISLPTSCASA 118
Qy 89 QNKDGPQEIIPVQSIHPCYNSSDVEDHNDLMLLQDQASLGSKVKPISLADHCTQP 148
Db 119 GTQCLISGWNKTSKSGTSYDVLKCLAPILSDSCSKSAYPGQITSNMFCAGYEGGKDS 178
Qy 149 GQKCTVSGWGTVPSPRENFPTLCAEVKIPQKCEDAPGQITDGMVAGSKGA-DT 207
Db 179 CQSDSGPVVCSGKLGQIVSGSG-CAQKNKPGYTVKCNVSVNIKOTIASN 229
Qy 208 CQSDSGPLVCDGALQGITSWGSDPCGRSDKPGVYTNICRYLDWIKKIIGSK 259

RESULT 7
ENTRY trypsin #type complete
TITLE trypsin (EC 3.4.21.4) precursor, cationic - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
#accession A27547
#authors Fletcher, T.S.; Alhadeff, M.; Craik, C.S.; Largman, C.
#journal Biochemistry (1987) 26:3081-3086
#title Isolation and characterization of a cDNA encoding rat cationic trypsinogen.
#cross-references MUID:8721609
#accession A27547
#molecule_type mRNA
#residues 1-247 #label FILE
#cross-references GB:M16624; NID:Q206498; PID:Q206499
CLASSIFICATION #superfamily trypsin; trypsin homology
KEYWORDS calcium binding; hydrolase; protein digestion; serine proteinase
FEATURE
25-240 #domain trypsin homology #label TRY\
31-161,49-65, #disulfide_bonds #status predicted\
133-234,140-207, #active_site His, Asp, Ser #status predicted\
172-186 #binding_site calcium (Glu, Asn, Val, Glu) #status
64,108,201 predicted
76,78,81,86
SUMMARY #length 247 #molecular-weight 26269 #checksum 9564
Query Match 40.3%; Score 783; DB 2; Length 247;
Best Local Similarity 43.0%; Pred. No. 7.53e-158;
Matches 108; Conservative 55; Mismatches 82; Indels 6; Gaps 6;
Db 2 KALIFLAFGLAAVALPLDDDDKIVGGYTCQKNSLPYQVSLNAGYHF-CGSLINQWVY 60
Qy 10 KTWFWLLGGWAGHAGRAQEDKVLGGHECOPHSQWQAALFOGQQLLGGVVGWVWL 69
Db 61 SAAHCYKSIQVRLGEHNDIVVEGGEQFIDAAKILIRPSYANFTF-D-N-DIMLIKLSNP 117
Qy 70 TAAHCKPKYTVRLGDHSLQNKDGPQEIIPVQSIHPCYNSSDVEDHNDLMLLQDQ 129
Db 118 ATLNSRVSTVSLPRCSGSGTKCLVSGWNTLSSGTWYPSLLQCLDAPVLSDSCKSSYP 177
Qy 130 ASLGSVKPISLADHCTQPQKCTVSGWGTVPSPRENFPTLCAEVKIPQKCEDAYP 189
Db 178 KGISNMFCLFLEGGKDSQCGSDSGPVVCGNQLQGVVSGYG-CAQKGPYTVKVCNY 236
Qy 190 QGITDGMVAGSKGA-DTCQSDGGLVCDGALQGITSWGSDPCGRSDKPGVYTNICRY 248
Db 237 VNWITQOTVAAN 247
Qy 249 LDWIKKIIGSK 259

```

```

RESULT 8
ENTRY trypsin #type complete
TITLE trypsin (EC 3.4.21.4) precursor (with pancreatic secretory

```

```

ENTRY trypsin #type complete
TITLE trypsin (EC 3.4.21.4) II precursor - rat
ALTERNATE_NAMES trypsinogen II
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 05-Apr-1983 #sequence_revision 30-Sep-1987 #text_change
18-Jul-1997
ACCESSIONS A22657; A00949
REFERENCE A22657
#authors Craik, C.S.; Choo, Q.L.; Swift, G.H.; Quinto, C.; MacDonald, R.J.; Rutter, W.J.
#journal J. Biol. Chem. (1984) 259:14255-14264
#title Structure of two related rat pancreatic trypsin genes.
#cross-references MUID:85054880
#accession A22657
#molecule_type DNA
#residues 1-246 #label CRA
REFERENCE A00948
#authors MacDonald, R.J.; Stary, S.J.; Swift, G.H.
#journal J. Biol. Chem. (1982) 257:9724-9732
#title Two similar but nonallelic rat pancreatic trypsinogens.
#cross-references MUID:82265624
#accession A00949
#molecule_type mRNA
#residues 9-246 #label MAC
COMMENT The trypsin II mRNA is present in much lower quantities than the trypsin I mRNA.
GENETICS 14/1; 67/2
#introns
CLASSIFICATION #superfamily trypsin; trypsin homology
KEYWORDS hydrolase; pancreas; protein digestion; serine proteinase; zymogen
FEATURE
1-15 #domain signal sequence #status predicted #label SIG\
16-23 #domain activation peptide #status predicted #label AP\
24-246 #product trypsin II #status predicted #label ENZ\
24-239 #domain trypsin homology #label TRY\
30-160,48-64, #disulfide_bonds #status predicted\
132-233,139-206, #active_site His, Asp, Ser #status predicted\
171-185 #binding_site calcium (Glu, Asn, Val, Glu) #status
63,107,200 predicted
75,77,80,85
SUMMARY #length 246 #molecular-weight 26243 #checksum 3816
Query Match 40.2%; Score 781; DB 1; Length 246;
Best Local Similarity 42.6%; Pred. No. 2.33e-157;
Matches 107; Conservative 56; Mismatches 81; Indels 7; Gaps 7;
Db 2 RALLFLALVGA-AVAPFVDDDDKIVGGYTCQENSVPYQVSLNSGYHF-CGSLINQWVY 59
Qy 10 KTWFWLLGGWAGHAGRAQEDKVLGGHECOPHSQWQAALFOGQQLLGGVVGWVWL 69
Db 60 SAAHCYKSIQVRLGEHNDIVVEGGEQFIDAAKILIRPNFDRKTL-N-N-DIMLIKLSNP 116
Qy 70 TAAHCKPKYTVRLGDHSLQNKDGPQEIIPVQSIHPCYNSSDVEDHNDLMLLQDQ 129
Db 117 VKLNARVAIVLPSGSCAPAGTQCLISGWNLTSSGYNEDLQCLDAPLPAQDCASYP 176
Qy 130 ASLGSVKPISLADHCTQPQKCTVSGWGTVPSPRENFPTLCAEVKIPQKCEDAYP 189
Db 177 KGITDNMVCVGLFEGGKDSQCGSDSGPVVCGNQLQGVVSGYG-CALPDNPGYTVKVCNY 235
Qy 190 QGITDGMVAGSKGA-DTCQSDGGLVCDGALQGITSWGSDPCGRSDKPGVYTNICRY 248
Db 236 VDWIQTIAAN 246
Qy 249 LDWIKKIIGSK 259

```

```

RESULT 9
ENTRY trypsin #type complete
TITLE trypsin (EC 3.4.21.4) precursor (with pancreatic secretory

```

```

ORGANISM      #trypsin inhibitor), chain Z - bovine
REFERENCE     #formal_name Bos primigenius taurus #common_name cattle
#note        pancreas
#authors     A50955
#submitters  Bolognesi, M.; Gatti, G.; Menegatti, E.; Guarneri, M.;
#cross-references submitted to the Brookhaven Protein Data Bank, September 1982
#cross-references PDB:1LGS
REFERENCE     TN005560
#authors     Bolognesi, M.; Gatti, G.; Menegatti, E.; Guarneri, M.;
#journal     Marquart, M.; Papamokos, E.; Huber, R.
#title       Three-dimensional structure of the complex between pancreatic
secretory inhibitor (kazal type) and trypsinogen at 1.8
angstroms resolution. structure solution, crystallographic
refinement and preliminary structural interpretation.
REFERENCE     TN005561
#authors     Marquart, M.; Walter, J.; Deisenhofer, J.; Bode, W.; Huber,
R.
#journal     Acta Crystallogr. (1983) B39:480
#title       The geometry of the reactive site and of the peptide groups
in trypsin, trypsinogen and its complexes with inhibitors.
REFERENCE     TN005562
#authors     Dayhoff, M.O.
#book        in Atlas of Protein Sequence and Structure (Data Section),
Dayhoff, M.O., ed., vol. 5, pp.105, National Biomedical
Research Foundation, Silver Spring, Md., 1972
REFERENCE     TN005563
#authors     Dayhoff, M.O.
#book        in Atlas of Protein Sequence and Structure, Supplement 1,
Dayhoff, M.O., ed., vol. 5, pp.88, National Biomedical
Research Foundation, Silver Spring, Md., 1973
COMMENT      Resolution: 1.8 angstroms
COMMENT      R-value: 0.186
COMMENT      Determination: X-ray diffraction
KEYWORDS     hydrolase; serine proteinase
FEATURE      146-158
#region helix (right hand alpha) (sngl alpha turn, rest
irreg.)\
210-215      #region helix (right hand 3-10) (contiguous with h3)\
216-225      #region helix (right hand alpha) (contiguous with h2)\
9-139        #disulfide_bonds\
27-43        #disulfide_bonds\
111-212      #disulfide_bonds\
118-185      #disulfide_bonds\
150-164      #disulfide_bonds\
175-199      #disulfide_bonds\
#length 225 #molecular-weight 23548 #checksum 5415
QUERY MATCH 40.1%; Score 779; DB 5; Length 225;
Best Local Similarity 47.0%; Pred. No. 7.22e-157;
Matches 108; Conservative 48; Mismatches 68; Indels 6; Gaps 6;
Db 1 DKIVGYTCGANTVPYQVLSNGYHF-CGGSLINSQWVYSAHCKYKGIQVRLGEDNINVE 59
Qy 31 DKVLGGHCQPHSQFQWQAALFQGLLGGVLVGNWVLTAAHCKPKYTVRLGDHSLQND 90
Db 60 VEGNEQFISAKSIIVHPSYNSNTL-N-DIMLIKLSAASLNSRVASISLPTSCASAGT 116
Qy 91 KDGPEQIIPVQSIHPHCYNSSDVEDHNDMLQLRDQASLGSVKRPISLADHCTOPGQ 150
Db 117 QCLISGWGNTKSSGTSPDYVLKCLKAPILSDSSCKSAYPGQITSNMFCAGYLEGGKDSQC 176
Qy 151 KCTVSGWGTVPSPRENFDTLNCAEVKIFPQKCEDAYPGQITDGMVCAGSKGA-DTCQ 209
Db 177 GDSGGPVVCSKLGIVSWGSG-CAQKNKPGVYTKVCNVYVSWIKOTIASN 225
Qy 210 GDSGGLVCDGALQGITSWGSDPCGRSDKPGVYTNICRYLDWIKKIIGSK 259
RESULT 10
ENTRY      2TLDE
TITLE      trypsin (EC 3.4.21.4) (with mutant streptomyces subtilisin

```

```

inhibitor), chain E - bovine
#formal_name Bos primigenius taurus #common_name cattle
REFERENCE     A50948
#authors     Mitsui, Y.; Takeuchi, Y.; Nonaka, T.; Nakamura, K.T.
#submitters  submitted to the Brookhaven Protein Data Bank, September 1991
#cross-references PDB:2TLD
REFERENCE     TN005550
#authors     Takeuchi, Y.; Nonaka, T.; Nakamura, K.T.; Kojima, S.; Miura,
K.I.; Mitsui, Y.
#journal     Proc. Natl. Acad. Sci. U.S.A. (1992) 89:4407
#title       Crystal structure of an engineered subtilisin inhibitor
complexed with bovine trypsin.
REFERENCE     TN005551
#authors     Takeuchi, Y.; Satow, Y.; Nakamura, K.T.; Mitsui, Y.
#journal     J. Mol. Biol. (1991) 221:309
#title       Refined crystal structure of the complex of subtilisin BPN'
and streptomyces subtilisin inhibitor at 1.8 angstroms
resolution.
COMMENT      Resolution: 2.6 angstroms
COMMENT      R-value: 0.173
COMMENT      Determination: X-ray diffraction
KEYWORDS     hydrolase; serine proteinase
FEATURE      26-28, 35-38, 85-89,
69-73        #region beta sheet\
46-50, 63-67 #region beta sheet\
136-141, 116-120,
179-183, 186-193,
204-209, 159-163,
144-152
#region beta sheet\
#region helix (right hand alpha) (single alpha turn, rest
irreg.)\
208-213      #region helix (right hand 3-10) (contiguous with h3)\
214-223      #region helix (right hand alpha) (contiguous with h2)\
7-137, 25-41,
109-210, 116-183,
148-162, 173-197,
84, 40, 177
#disulfide_bonds\
#active_site Asp, His, Ser #label CAT\
#site #label S13
#length 223 #molecular-weight 23071 #checksum 5569
SUMMARY
Query Match 39.5%; Score 768; DB 5; Length 223;
Best Local Similarity 46.5%; Pred. No. 3.60e-154;
Matches 106; Conservative 48; Mismatches 68; Indels 6; Gaps 6;
Db 1 IVGYTCGANTVPYQVLSNGYHF-CGGSLINSQWVYSAHCKYKGIQVRLGEDNINVE 59
Qy 33 VLGGHCQPHSQFQWQAALFQGLLGGVLVGNWVLTAAHCKPKYTVRLGDHSLQND 92
Db 60 GNEQFISAKSIIVHPSYNSNTL-N-N-DIMLIKLSAASLNSRVASISLPTSCASAGTQC 116
Qy 93 GPEQIIPVQSIHPHCYNSSDVEDHNDMLQLRDQASLGSVKRPISLADHCTOPGQKC 152
Db 117 LISGWGNTKSSGTSPDYVLKCLKAPILSDSSCKSAYPGQITSNMFCAGLEGGDSQCGD 176
Qy 153 TVSGWGTVPSPRENFDTLNCAEVKIFPQKCEDAYPGQITDGMVCAGS-SKGADTCQGD 211
Db 177 SGPVTVCSKLGIVSWGSG-CAAKNKPVTYTKVCNVYVSWIKOTIASN 223
Qy 212 SGGPLVCDGALQGITSWGSDPCGRSDKPGVYTNICRYLDWIKKIIGSK 259
RESULT 11
ENTRY      B25528
TITLE      trypsin (EC 3.4.21.4) precursor - mouse
ORGANISM     #formal_name Mus musculus #common_name house mouse
DATE        30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change
20-Mar-1998
ACCESSIONS B25528
REFERENCE     A93646
#authors     Stevenson, B.J.; Hagenbuechle, O.; Wellauer, P.K.
#journal     Nucleic Acids Res. (1986) 14:8307-8330
#title       Sequence organisation and transcriptional regulation of the
mouse elastase II and trypsin genes.

```

```

#cross-references MUID:87066713
#accession B25528
#molecule_type mRNA
#residues 1-246 #label STE
#cross-references GB:X04574; NID:g54918; PID:g54919
CLASSIFICATION #superfamily trypsin; trypsin homology
KEYWORDS calcium binding; hydrolase; protein digestion; serine
proteinase
FEATURE
1-23 #domain signal sequence #status predicted #label SIG\
24-246 #product trypsin #status predicted #label MAT\
24-239 #domain trypsin homology #label TRY\
30-160,48-64, #disulfide_bonds #status predicted\
132-233,139-206, #active_site His, Asp, Ser #status predicted\
171-185 #binding_site calcium (Glu, Asn, Val, Glu) #status
63,107,200 predicted
75,77,80,85
SUMMARY #length 246 #molecular-weight 26203 #checksum 5833
Query Match 39.5%; Score 768; DB 2; Length 246;
Best Local Similarity 42.7%; Pred. No. 3,60e-154;
Matches 106; Conservative 56; Mismatches 79; Indels 7; Gaps 7;
Db 5 L1ALVGAA-VAFPVDDDKIVGGTCRESSVPYQVSLNAGYHF-CGSLINDQWVYSA 62
QY 13 MFLLLGGAWAGHSRAQEDKVLGHECQPHSQPQAAALFQGLCGVVGNNVLTAA 72
Db 63 HCYKRTQVRLGHNINVLGNEQFVDSAKIIRHPNNSWTL-D-N-DIMLIKASPVTL 119
QY 73 HCKPKYTVRLGHSLSQNKDQPEIPVQSPHPHCYNSSDVEDHNDMLLQLRQASL 132
Db 120 NARVAPVLPSSCAPAGTCOLISGNGNTLSNGVNPDLQCVDPVLPQADCEASYPGDI 179
QY 133 GSKVKPISLADHCTQPGQKCTVSGWGTVPSPRENFDTLNCAEVKIFPKKCEDAIPGQI 192
Db 180 TNNMICVGFLEGGKDCQSGSGGPPVNCGLQIVSMGYG-CAQDPAGVYTKVNTVDW 238
QY 193 TDGMVACAGSSKGA-DTCQGDGSGPLVCDGALQGITSGSDPCGRSDRPGVYTNICRYLDW 251
Db 239 IQNTIADN 246
QY 252 IKKIIGSK 259

RESULT 12 TRRT1 #type complete
ENTRY trypsin (EC 3.4.21.4) I precursor - rat
TITLE trypsinogen I
FORMAL_NAME Rattus norvegicus #common_name Norway rat
REFERENCE 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change
20-Mar-1998
ACCESSIONS B22657; A00948
REFERENCE A22657
#authors Craik, C.S.; Choo, Q.L.; Swift, G.H.; Quinto, C.; MacDonald,
R.J.; Rutter, W.J.
#journal J. Biol. Chem. (1984) 259:14255-14264
#cross-references MUID:85054880
#accession B22657
#molecule_type DNA
#residues 1-246 #label CRA
#cross-references GB:J00778; NID:g206507; PID:g206508
#note the authors translated the codon ATC for residue 6 as
Leu and GAC for residue 170 as Asn
REFERENCE A00948
#authors MacDonald, R.J.; Stary, S.J.; Swift, G.H.
#journal J. Biol. Chem. (1984) 257:9724-9732
#title Two similar but nonallelic rat pancreatic trypsinogens.
Nucleotide sequences of the cloned cDNAs.
#cross-references MUID:82265624
#accession A00948
#molecule_type mRNA

```

```

#residues 1-246 #label MAC
GENETICS
#introns 14/1: 67/2: 152/1: 197/3
CLASSIFICATION #superfamily trypsin; trypsin homology
KEYWORDS hydrolase; pancreas; protein digestion; serine proteinase;
zymogen
FEATURE
1-15 #domain signal sequence #status predicted #label SIG\
16-23 #domain activation peptide #status predicted #label ACT\
24-246 #product trypsin I #status predicted #label ENZ\
24-239 #domain trypsin homology #label TRY\
30-160,48-64, #disulfide_bonds #status predicted\
132-233,139-206, #active_site His, Asp, Ser #status predicted\
171-185 #binding_site calcium (Glu, Asn, Val, Glu) #status
63,107,200 predicted
75,77,80,85
SUMMARY #length 246 #molecular-weight 25959 #checksum 6732
Query Match 39.5%; Score 767; DB 1; Length 246;
Best Local Similarity 42.3%; Pred. No. 6,33e-154;
Matches 105; Conservative 56; Mismatches 80; Indels 7; Gaps 7;
Db 5 L1ALVGAA-VAPFLEDDDKIVGGTCPEHSVPYQVSLNAGYHF-CGSLINDQWVYSA 62
QY 13 MFLLLGGAWAGHSRAQEDKVLGHECQPHSQPQAAALFQGLCGVVGNNVLTAA 72
Db 63 HCYKRTQVRLGHNINVLGDEQFVNAKIIKHPNYSWTL-N-N-DIMLIKASPVKL 119
QY 73 HCKPKYTVRLGHSLSQNKDQPEIPVQSPHPHCYNSSDVEDHNDMLLQLRQASL 132
Db 120 NARVAPVLPSSCAPAGTCOLISGNGNTLSNGVNPDLQCVDPVLPQADCEAAVPGEI 179
QY 133 GSKVKPISLADHCTQPGQKCTVSGWGTVPSPRENFDTLNCAEVKIFPKKCEDAIPGQI 192
Db 180 TNNMICVGFLEGGKDCQSGSGGPPVNCGLQIVSMGYG-CALPONPGYTKVNCVFGW 238
QY 193 TDGMVACAGSSKGA-DTCQGDGSGPLVCDGALQGITSGSDPCGRSDRPGVYTNICRYLDW 251
Db 239 IQNTIADN 246
QY 252 IKKIIGSK 259

RESULT 13 TRPS #type complete
ENTRY Trypsin complexed with inhibitor a90720a - bovine
TITLE #formal_name Bos primigenius taurus #common_name cattle
#note pancreas
REFERENCE A67200
#authors Lee, A.Y.; Clardy, J.
#submission submitted to the Brookhaven Protein Data Bank, September 1994
#cross-references PDB:1TPS
REFERENCE TN023754
#authors Lee, A.Y.; Smitka, T.A.; Bonjouklian, R.; Clardy, J.
#journal Chem. Biol. (1994) 1:113
#title Atomic structure of the trypsin-a90720a complex: a unified
approach to structure and function.
COMMENT Resolution: 1.9 angstroms
DETERMINATION: X-ray diffraction
KEYWORDS Hydrolase(serine protease)
FEATURE
144-156 #region helix (right hand alpha) (sngl alpha turn, rest
irreg.)\
208-213 #region helix (right hand 3-10) (contiguous with h3)\
214-223 #region helix (right hand alpha) (contiguous with h2)\
7-137 #disulfide_bonds\
25-41 #disulfide_bonds\
109-210 #disulfide_bonds\
116-183 #disulfide_bonds\
148-162 #disulfide_bonds\
173-197 #disulfide_bonds\
SUMMARY #length 223 #molecular-weight 23305 #checksum 7159

```

Query Match 39.4%; Score 766; DB 5; Length 223;
Best Local Similarity 46.5%; Pred. No. 1.11e-153;
Matches 106; Conservative 48; Mismatches 68; Indels 6; Gaps 6;

Db 1 IVGGYTCGANTVPYQVSLNSGYHF-CGGSLSINSOWVYSAACHYKSGIQVRLGEDNINWVE 59
QY 33 VLGHECQPHSQPQWQALFQCGVLLGGVNGVWVTAACHKKPKYTVRLGDHSLQNKD 92

Db 60 GNEQFISAKSIVHPSYNSNTL-N-DIMLIKLSAASLSNRVASISLPTSCASAGTQC 116
QY 93 GPEQEIIPVQSIHPHCYNSDVEDHNDMLLQLRDQASLGSKVKPISLADHCTQPGQKC 152

Db 117 LISGWNKTSSTGYPDVVKLKLAPILSDSCSKSAYPQGITSNMFCAGYLEGGKDCSQGD 176
QY 153 TVSGWGTVTSPRENFPTLNCAYKIFPQKKCEDAYPQGITDGMVCAGSSKGA-DTCQGD 211

Db 177 SGGPVVCSGKLOGIVSWGSG-CAQKNKPGVYTKVNCVSVWIKOTIASN 223
QY 212 SGGPLVCDGALQGITSWGSDPCGRSDKPGVYTNICRYLDWIKKIIGSK 259

Search completed: Wed Jul 22 15:28:54 1998
Job time : 73 secs.

Query Match 39.4%; Score 766; DB 5; Length 223;
Best Local Similarity 46.5%; Pred. No. 1.11e-153;
Matches 106; Conservative 48; Mismatches 68; Indels 6; Gaps 6;

Db 1 IVGGYTCGANTVPYQVSLNSGYHF-CGGSLSINSOWVYSAACHYKSGIQVRLGEDNINWVE 59
QY 33 VLGHECQPHSQPQWQALFQCGVLLGGVNGVWVTAACHKKPKYTVRLGDHSLQNKD 92

Db 60 GNEQFISAKSIVHPSYNSNTL-N-DIMLIKLSAASLSNRVASISLPTSCASAGTQC 116
QY 93 GPEQEIIPVQSIHPHCYNSDVEDHNDMLLQLRDQASLGSKVKPISLADHCTQPGQKC 152

Db 117 LISGWNKTSSTGYPDVVKLKLAPILSDSCSKSAYPQGITSNMFCAGYLEGGKDCSQGD 176
QY 153 TVSGWGTVTSPRENFPTLNCAYKIFPQKKCEDAYPQGITDGMVCAGSSKGA-DTCQGD 211

Db 177 SGGPVVCSGKLOGIVSWGSG-CAQKNKPGVYTKVNCVSVWIKOTIASN 223
QY 212 SGGPLVCDGALQGITSWGSDPCGRSDKPGVYTNICRYLDWIKKIIGSK 259

Search completed: Wed Jul 22 15:28:54 1998
Job time : 73 secs.

RESULT 15

ENTRY 1JRTA #type complete
TITLE trypsin (EC 3.4.21.4), chain A - bovine
PDB TITLE hemiacetal complex between leupeptin and trypsin
ORGANISM #formal_name Bos taurus; #common_name bovine
REFERENCE A65928
#authors Kurinov, I.V.; Harrison, R.W.
#submission submitted to the Brookhaven Protein Data Bank, February 1996
#cross-references PDB:1JRT
REFERENCE TN027917
#authors Kurinov, I.V.; Harrison, R.W.
#journal Nat. Struct. Biol. (1994) 1:735
#title Prediction of new serine proteinase inhibitors.
COMMENT Resolution: 1.8 angstroms
COMMENT Determination: X-ray diffraction
COMMENT R-value: no refinement
KEYWORDS digestion; hydrolase; pancreas; serine protease; zymogen
FEATURE
39-41 #region helix (right hand 3-10)\
145-151 #region helix (right hand alpha)\
209-211 #region helix (right hand 3-10)\
23-31, 34-37, 86-90, #region beta sheet\
67-72, 136-141, 115-120, #region beta sheet\
136-141, 115-120, #region beta sheet\
180-183, 186-193, #disulfide_bonds\
204-208, 160-163, #disulfide_bonds\
7-137 #disulfide_bonds\
25-41 #disulfide_bonds\
109-210 #disulfide_bonds\
116-183 #disulfide_bonds\
148-162 #disulfide_bonds\
173-197 #disulfide_bonds\
SUMMARY #length 223 #molecular-weight 23305 #checksum 7159

Query Match 39.4%; Score 766; DB 5; Length 223;
Best Local Similarity 46.5%; Pred. No. 1.11e-153;
Matches 106; Conservative 48; Mismatches 68; Indels 6; Gaps 6;

Db 1 IVGGYTCGANTVPYQVSLNSGYHF-CGGSLSINSOWVYSAACHYKSGIQVRLGEDNINWVE 59
QY 33 VLGHECQPHSQPQWQALFQCGVLLGGVNGVWVTAACHKKPKYTVRLGDHSLQNKD 92

Db 60 GNEQFISAKSIVHPSYNSNTL-N-DIMLIKLSAASLSNRVASISLPTSCASAGTQC 116
QY 93 GPEQEIIPVQSIHPHCYNSDVEDHNDMLLQLRDQASLGSKVKPISLADHCTQPGQKC 152

Db 117 LISGWNKTSSTGYPDVVKLKLAPILSDSCSKSAYPQGITSNMFCAGYLEGGKDCSQGD 176
QY 153 TVSGWGTVTSPRENFPTLNCAYKIFPQKKCEDAYPQGITDGMVCAGSSKGA-DTCQGD 211

Db 177 SGGPVVCSGKLOGIVSWGSG-CAQKNKPGVYTKVNCVSVWIKOTIASN 223
QY 212 SGGPLVCDGALQGITSWGSDPCGRSDKPGVYTNICRYLDWIKKIIGSK 259

Search completed: Wed Jul 22 15:28:54 1998
Job time : 73 secs.

THIS PAGE BLANK (USPTO)

MPARLH (TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MParch_pp protein - protein database search, using Smith-Waterman algorithm

on: Wed Jul 22 15:29:54 1998; MasPar time 15.26 Seconds
Tabular output not generated. 717.292 Million cell updates/sec

Title: >US-08-915-659A-7
Description: (1-260) from US08915659A.pep
Perfect Score: 1943
Sequence: 1 MGRPRPRAAKTWFWLLLG.....VTNCRYLWDWIKKIGSKG 260

Scoring table: PAM 150
Gap 11

Searched: 140542 seqs, 42109429 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: spiremb15
1:sp_fungi 2:sp_human 3:sp_invertebrate 4:sp_mammal
5:sp_mmc 6:sp_organelle 7:sp_phage 8:sp_plant
9:sp_bacteria 10:sp_rodent 11:sp_virus 12:sp_vertebrate
13:sp_unclassified

Statistics: Mean 42.195; Variance 59.229; scale 0.712

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1488	76.6	260	10	SERINE PROTEASE INHIBI	0.00e+00
2	843	43.4	244	2	PROTEASE M PRECURSOR.	1.15e-206
3	776	39.9	273	2	NS1-NORMAL EPITHELIAL	5.59e-187
4	757	39.0	242	12	TRYPsin PRECURSOR (EC	2.06e-181
5	754	38.8	237	12	TRYPsin (FRAGMENT)	1.55e-180
6	751	38.7	247	12	TRYPsin A2 PRECURS	1.17e-179
7	751	38.7	261	10	TISSUE KALLIKREIN.	1.17e-179
8	750	38.6	247	12	TRYPsin A3 PRECURS	2.30e-179
9	744	38.3	245	12	TRYPsin B2 PRECURS	1.31e-177
10	738	38.0	244	12	TRYPsin B1 PRECURS	7.44e-176
11	710	36.5	259	2	TRYPsin IV B-FORM.	1.12e-167
12	703	36.2	258	4	KALLIKREIN PRECURSOR.	1.23e-165
13	703	36.2	261	4	KALLIKREIN PRECURSOR (1.23e-165
14	694	35.7	239	10	KALLIKREIN (FRAGMENT).	5.14e-163
15	663	34.1	249	12	PREPROTRYPsin PRECURSO	5.33e-154
16	659	33.9	235	10	KALLIKREIN (FRAGMENT).	7.73e-153
17	644	33.1	219	12	TRYPsin I (FRAGMENT)	1.73e-148
18	613	31.5	258	12	VENOM PLASMINOGEN ACTI	1.62e-139
19	593	30.5	262	12	CALOSIN	9.53e-134
20	590	30.4	260	12	SERINE PROTEASE PRECUR	6.97e-133

21	581	29.9	258	12	013059	SERINE PROTEASE PRECUR	2.72e-130
22	579	29.8	260	12	013056	SERINE PROTEASE PRECUR	1.02e-129
23	576	29.6	250	10	003955	GLANDULAR KALLIKREIN (7.46e-129
24	570	29.3	258	12	042207	CAPILLARY PERMEABILITY	3.96e-127
25	568	29.2	258	12	013063	SERINE PROTEASE PRECUR	1.49e-126
26	565	29.1	257	12	091509	MUCOFIRASE 3.	1.08e-125
27	560	28.8	260	12	013057	SERINE PROTEASE PRECUR	2.94e-124
28	557	28.7	257	12	013062	SERINE PROTEASE PRECUR	2.13e-123
29	552	28.4	257	12	091507	MUCOFIRASE 1.	5.79e-122
30	545	28.0	257	12	091508	MUCOFIRASE 2.	5.86e-120
31	545	28.0	257	12	091500	PREPROTRIMUBIN PRECURS	5.86e-120
32	538	27.7	257	12	091510	MUCOFIRASE 4.	5.90e-118
33	537	27.6	257	12	091511	MUCOFIRASE 5.	1.14e-117
34	520	26.8	257	12	013058	SERINE PROTEASE PRECUR	8.16e-113
35	521	26.8	263	10	035205	GRANZYME K.	4.23e-113
36	518	26.7	257	12	013069	KN-BJ2 PRECURSOR.	3.03e-112
37	509	26.2	258	10	061280	PRECURSOR ADIPSIN.	1.11e-109
38	504	25.9	248	10	063636	NATURAL KILLER CELL PR	2.94e-108
39	502	25.8	260	12	013061	SERINE PROTEASE PRECUR	1.09e-107
40	480	24.7	248	10	063224	GRANZYME-LIKE PROTEIN	1.91e-101
41	476	24.5	264	10	008643	LYMPHOCTYE MET-ASE 1 P	2.58e-100
42	476	24.5	271	3	018487	CHYMOTRYPSIN BI (FRAGM	2.58e-100
43	473	24.3	135	10	Q62284	NERVE GROWTH FACTOR, G	1.82e-99
44	471	24.2	225	2	Q15358	SERINE PROTEASE HOMOLO	6.71e-99
45	459	23.6	560	2	Q14520	HGF ACTIVATOR LIKE PRO	1.64e-95

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	260 AA.
AC	Q61955			
DT	01-NOV-1996	(TREMREL. 01, CREATED)		
DT	01-NOV-1996	(TREMREL. 01, LAST SEQUENCE UPDATE)		
DT	01-JAN-1998	(TREMREL. 05, LAST ANNOTATION UPDATE)		
DE	SERINE PROTEASE INHIBITOR 5 (NEUROPSIN).			
GN	SP15			
OS	MUS MUSCULUS (MOUSE).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; RODENTIA.			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RC	STRAIN-BALB/C; TISSUE-HIPPOCAMPUS;			
RX	MEDLINE; 95348817.			
RA	ZU-LIN C., YOSHIDA S., KATO K., MOMOTA Y., SUZUKI J., TANAKA T.,			
RA	ITO J., NISHINO H., AIMOTO S., KIYAMA H.;			
RL	J. NEUROSCI. 15:5088-5097(1995).			
DR	EMBL; D30785; G1020091; -.			
DR	MGI; 104519; SP15.			
KW	SERINE PROTEASE.			
SQ	SEQUENCE 260 AA; 28523 MW; 232F78DB CRC32;			
Query Match	76.6%;	Score 1488;	DB 10;	Length 260;
Best Local Similarity	72.2%;	Pred. No. 0.00e+00;		
Matches	187;	Conservative 36;	Mismatches 34;	Indels 0; Gaps 0;
Db	1	MGRPPPCAQTQPIILLIFMGAWAGLTRAQSKILGREGICPHSQPQWQAALFOGRLICGG	60	
Qy	1	MGRPRPRAAKTWFWLLLGAGNAGHRAQEDKVLGHECQPHSQPQWQAALFOGQQLICGG	60	
Db	61	VLVGRVWLVTAACHKQKQKQSVRLGDSHLSQSRQPEQIQVQAOISQHPCYNNNSDPDHS	120	
Qy	61	VLVGVNVLWLVTAACHKPKYTVRLGDSHLSQNRQPEQIPVQSQIPHPCYNSDDVDHND	120	
Db	121	IMLRLONSANLGDVKKVQLANLCPKVGOKCIISGWCVTTSPOENFNTLCAEVKIYS	180	
Qy	121	LMELQLRDQASLGSKVKPKYISLADHCTQPGQKCTVSGWGTVTSPRENFPTLCAEVKIFP	180	
Db	181	QNKCBRAYPGKITEGMVCAAGSNGADTCOGDSGGPLVCDMLQGITSWGSDPCGKPEKG	240	
Qy	181	QKCEDATPGQITDGMVCAAGSNGADTCOGDSGGPLVCDMLQGITSWGSDPCGSKDKFG	240	
Db	241	VYTKICRYTTWIKTKMDNR	259	

Db 232 NDWL 235
Qy 249 LDWI 252

RESULT 5

ID Q91515 PRELIMINARY; PRT; 237 AA.
AC Q91515;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE TRYPsinogen (FRAGMENT).
OS FUGU RUBRIPES (JAPANESE PUFFERFISH).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC OSTEICHTHYES; ACTINOPTERYGII; TETRAODONTIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RA WANG K., GAN L., LEE I., ROACH J., HOOD L.;
SUBMITTED (APR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
EMBL; U25747; G971196; -.
FT NON_TER 1
SQ SEQUENCE 237 AA; 25726 MW; 6A296989 CRC32;

Query Match 38.8%; Score 754; DB 12; Length 237;
Best Local Similarity 42.4%; Pred. No. 1.55e-180;
Matches 103; Conservative 64; Mismatches 68; Indels 8; Gaps 8;

Db 1 LIAAAYAA-PIDEDDKIVGVECRKNSVAYQVSLNSGYHF-CGSLYNNVNVSAHCYK 58
Qy 17 LGGAWAGHSRAQEDKVLGGHECQPHSQPWAALFQGGQLCGGVLVGNWLVTAACHCK 76
Db 59 SRVVYRLGEHNRANEGPEQFISSSRVIRHPNYSYNT-D-N-DIMLIKLSKATLQYV 115
Qy 77 PKYTVRLGDHSLQNKDGPQEIPIVQSPHPFCYNSDVEDHNDMLLQLRDQASLSKV 136
Db 116 OPVALPSSCAAAGTKVYSGNGTMSSTAD-RNKLCNLNIPILSDRCENSYPGMITDAM 174
Qy 137 KPISLADHCTQPGKCTVSGNGTVPSPRENFPTLNCAEVKIFFQKKCEDAYPGQITDGM 196
Db 175 FCAGYLEGKDSQCGDGGPVVCCNQLGVVSWGYG-CAERDHPGVYAKVCLFNWLEST 233
Qy 197 VCAGSSKGA-DTCQDGGPLVCDGALQGITSWGSDPCGRSKDPGVYTNICRYLDWIKKI 255

Db 234 MAS 236

Qy 256 IGS 258

RESULT 6

ID O42158 PRELIMINARY; PRT; 247 AA.
AC O42158;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DE TRYPsinogen A2 PRECURSOR.
GN TRYP2.
OS PETROMYZON MARINUS (SEA LAMPREY).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES;
OC AGNATHA (CYCLOSTOMATA).
RN [1]
RP SEQUENCE FROM N.A.
RA ROACH J.C.;
SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
EMBL; AF011898; G2367495; -.
DR EMBL.
KW SIGNAL.
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 16 247 TRYPsin A2.
SQ SEQUENCE 247 AA; 26309 MW; D049AA4C CRC32;

Query Match 38.7%; Score 751; DB 12; Length 247;
Best Local Similarity 42.0%; Pred. No. 1.17e-179;
Matches 105; Conservative 59; Mismatches 77; Indels 9; Gaps 9;

Db 4 LIALLVGVAAPYMYEDHIVGSECAHSPQWQVSLNIGYHF-CGSLINSQWVYSA 62
Qy 13 MFLILLGGAWAGHSRAQEDKVLGGHECQPHSQPWAALFQGGQLCGGVLVGNWLVTA 72
Db 63 HCYQTASRISVRIGEHNFVNEGTEQIOAKAIOHPQYNSWTI-D-N-DIMLIKLSPP 119
Qy 73 HC-KKP-KYTVRLGDHSLQNKDGPQEIPIVQSPHPFCYNSDVEDHNDMLLQLRDQA 130
Db 120 TLNOYADAIAIPSSCVNTGVMCTISGGETQTSIGS-PDVLQVQAPVLSDTSCRNYPG 178
Qy 131 SLGSKVAPISLADHCTQPGKCTVSGNGTVPSPRENFPTLNCAEVKIFFQKKCEDAYPG 190
Db 179 DITNNMICLGLGEGKDSQCGDGGPVVCCNQLGVVSWGRG-CALPNYPGVYTKVCNIN 237
Qy 191 QITDGMVCASSKGA-DTCQDGGPLVCDGALQGITSWGSDPCGRSKDPGVYTNICRYL 249
Db 238 AWTAQTAAAN 247
Qy 250 DWIKKIIGSK 259

RESULT 7

ID Q61855 PRELIMINARY; PRT; 261 AA.
AC Q61855;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE TISSUE KALLIKREIN.
GN MKR-6.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RA TADA M., PETERS J., TAKAHASHI S., INOUE H., MIYAKE Y.;
RL SUBMITTED (AUG-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; D10464; G520591; -.
SQ SEQUENCE 261 AA; 28775 MW; 509D1D00 CRC32;

Query Match 38.7%; Score 751; DB 10; Length 261;
Best Local Similarity 42.4%; Pred. No. 1.17e-179;
Matches 108; Conservative 57; Mismatches 78; Indels 12; Gaps 9;

Db 6 LFLALSUGGI-DAAPPVQ-SRIVGGENCKNSOPQWAVYRFTKYQCGGILLNANWVITA 63
Qy 13 MFLILL-LGGAWAGHSRAQEDKVLGGHECQPHSQPWAALFQGGQLCGGVLVGNWLVTA 71
Db 64 AHCHNDKYQVWLKKNFLEDEPSAQHRLVSKAIPHPDFNMNLSLNEHTPOPEDDYSNDLML 123
Qy 72 AHCKKPKYTVRLGDHSLQNKDGPQEIPIVQSPHPFCYN-S--SD-V--EDH-NHDLML 123
Db 124 LRLKPPADITDVVKPIDLPTEEPKLGSTCLASGWSITPVKYEYPDELOCVNKLKLPNE 183
Qy 124 LQLRDQASLGSKVKPISLADHCTQPGKCTVSGNGTVPSPRENFPTLNCAEVKIFFQKK 183
Db 184 CAAHIEKVTDDMLCAGDMGDKTACAGDSGGLICDGLQGITSWGSPCGKPNVPGIY 243
Qy 184 CEDAYPGQITDGMVCASSKGA-DTCQDGGPLVCDGALQGITSWGSDPCGRSKDPGVY 242
Db 244 TRVLFNTWIRETMA 258
Qy 243 TNICRYLDWIKKIIG 257

RESULT 8

ID O42608 PRELIMINARY; PRT; 247 AA.
AC O42608;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DE TRYPsinogen A3 PRECURSOR.
GN TRYP3.

OS PETROMYZON MARINUS (SEA LAMPREY).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES;
 OC AGNATHA (CYCLOSTOMATA).
 RN [1]
 RP SEQUENCE FROM N.A.
 RA ROACH J.C.;
 RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ANTERIOR INTESTINE;
 RA ROACH J.C.;
 RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; AF011899; G2367497; -;
 DR EMBL; AF011352; G2293478; -;
 KW SIGNAL.
 FT SIGNAL
 FT CHAIN 1 15 POTENTIAL.
 FT CHAIN 16 247 TRYPSIN A3.
 SQ SEQUENCE 247 AA; 26295 MW; 6D71AC2E CRC32;

Query Match 38.6%; Score 750; DB 12; Length 247;
 Best Local Similarity 42.0%; Pred. No. 2.30e-179;
 Matches 105; Conservative 59; Mismatches 77; Indels 9; Gaps 9;
 4 LILALLVGAAPYEDHIVGSECAHSPQWVSNIGYHF-CGGSLSINQWVWSAA 62
 QY 13 MFLLLGGAWAGHRAQEDKVLGGHECOPHSQPQWQALFQOQLCGVLGGNVLTA 72
 Db 63 HCQTASRISVRIGEHNFVTEGTEQRIQASKAIRHPQYNSATI-D-N-DIMLIKLSPPA 119
 QY 73 HC-KKP-KYTVRLGDHSLQNKDGPQEIIPVQSIHPHCYNSDVEDHNDMLQLRDQA 130
 Db 120 TLNQYQAQVLPSSCVGTGMCTISGGETQTSVGS-PDVLQVQAPVLSDTSCRSNYPG 178
 QY 131 SLGSKVKPISLADHCTOPGQKCTVSGWGTVPSPRENFDTLNCARVIFPQKKCEDAYPG 190
 Db 179 DITNNMICLGLGEGKDCOGSGPVCNGELQGVISWGRG-CALPNYPGVYTKVCNYN 237
 QY 191 QITDGMVCAAGSSKA-DTCQDGGPLVCDGALQGITSWGSDPCGRSKDPGVYTNICRYL 249
 Db 238 AWIAQTIAAN 247
 QY 250 DWIKKIIGSK 259

RESULT 9
 ID 042160 PRELIMINARY; PRT; 245 AA.
 AC 042160;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DE TRYPB2.
 GN TRYPB2.
 OS PETROMYZON MARINUS (SEA LAMPREY).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES;
 OC AGNATHA (CYCLOSTOMATA).
 RN [1]
 RP SEQUENCE FROM N.A.
 RA ROACH J.C.;
 RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; AF011901; G2367501; -;
 KW SIGNAL.
 FT SIGNAL
 FT CHAIN 1 1 POTENTIAL.
 FT CHAIN 14 245 TRYPSIN B2.
 SQ SEQUENCE 245 AA; 26001 MW; 071872F0 CRC32;

Query Match 38.3%; Score 744; DB 12; Length 245;
 Best Local Similarity 42.0%; Pred. No. 1.31e-177;
 Matches 105; Conservative 61; Mismatches 74; Indels 10; Gaps 10;
 3 IFALLVGTGA-AAAPYMWEDHIVGSECAHSPQWVSNIGYHF-CGGSLSISSEWVWSAA 60
 QY 13 MFLLLGGAWAGHRAQEDKVLGGHECOPHSQPQWQALFQOQLCGVLGGNVLTA 72

Db 61 HCQTASRISVRIGEHNFVTEGTEQRIQASKAIRHPQYNSATI-D-N-DIMLIKLSPPA 117
 QY 73 HC-KKP-KYTVRLGDHSLQNKDGPQEIIPVQSIHPHCYNSDVEDHNDMLQLRDQA 130
 Db 118 TLNQYQAQVLPSSCVGTGMCTISGGETQTSVGS-PDVLQVQAPVLSDTSCRSNYPG 176
 QY 131 SLGSKVKPISLADHCTOPGQKCTVSGWGTVPSPRENFDTLNCARVIFPQKKCEDAYPG 190
 Db 177 DITNNMICLGLGEGKDCOGSGPVCNGELQGVISWGRG-CALPNYPGVYTKVCNYN 235
 QY 191 QITDGMVCAAGSSKA-DTCQDGGPLVCDGALQGITSWGSDPCGRSKDPGVYTNICRYL 249
 Db 236 SWIASTMAAN 245
 QY 250 DWIKKIIGSK 259

RESULT 10
 ID 042159 PRELIMINARY; PRT; 244 AA.
 AC 042159;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DE TRYPB1.
 GN TRYPB1.
 OS PETROMYZON MARINUS (SEA LAMPREY).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES;
 OC AGNATHA (CYCLOSTOMATA).
 RN [1]
 RP SEQUENCE FROM N.A.
 RA ROACH J.C.;
 RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; AF011900; G2367499; -;
 KW SIGNAL.
 FT SIGNAL
 FT CHAIN 1 1 NON_TER
 FT CHAIN 13 12 POTENTIAL.
 FT CHAIN 13 244 TRYPSIN B1.
 SQ SEQUENCE 244 AA; 25903 MW; 41F78768 CRC32;

Query Match 38.0%; Score 738; DB 12; Length 244;
 Best Local Similarity 41.2%; Pred. No. 7.44e-176;
 Matches 103; Conservative 63; Mismatches 74; Indels 10; Gaps 10;

Db 2 IFALLVGTGA-AAAPYMWEDHIVGSECAHSPQWVSNIGYHF-CGGSLSISSEWVWSAA 59
 QY 13 MFLLLGGAWAGHRAQEDKVLGGHECOPHSQPQWQALFQOQLCGVLGGNVLTA 72
 Db 60 HCQTASRISVRIGEHNFVTEGTEQRIQASKAIRHPQYNSATI-D-N-DIMLIKLSPPA 116
 QY 73 HC-KKP-KYTVRLGDHSLQNKDGPQEIIPVQSIHPHCYNSDVEDHNDMLQLRDQA 130
 Db 117 TLNQYQAQVLPSSCVGTGMCTISGGETQTSVGS-PDVLQVQAPVLSDTSCRSNYPG 175
 QY 131 SLGSKVKPISLADHCTOPGQKCTVSGWGTVPSPRENFDTLNCARVIFPQKKCEDAYPG 190
 Db 176 DITNNMICLGLGEGKDCOGSGPVCNGELQGVISWGRG-CALPNYPGVYTKVCNYN 234
 QY 191 QITDGMVCAAGSSKA-DTCQDGGPLVCDGALQGITSWGSDPCGRSKDPGVYTNICRYL 249
 Db 235 SWIASTMAAN 244
 QY 250 DWIKKIIGSK 259

RESULT 11
 ID Q15655 PRELIMINARY; PRT; 259 AA.
 AC Q15655;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DE TRYPB1.
 GN TRYPB1.
 OS HOMO SAPIENS (HUMAN).

RA ZINTZ C.B., MA J.X., CHAO J., CHAO L.;
 RL SUBMITTED (JAN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; L33840; G609585; -;
 FT NON_TER 1
 SQ SEQUENCE 239 AA; 26382 MW; 0609E3E6 CRC32;

Query Match 35.7%; Score 694; DB 10; Length 239;
 Best Local Similarity 44.9%; Pred. No. 5.14e-163;
 Matches 106; Conservative 52; Mismatches 63; Indels 15; Gaps 11;

Db 6 QSRVVGFEKCEKNSQWQVAVIN-EDL-CGGVLDPSWVITAACHYSDNYHVLGQNNL- 62
 Qy 30 EDKVLGGHECPHSPQWQAALFQGGVLLCGVLYGGNWLTAACHCKPKYTVRLGDHSLQ 89
 Db 63 SEDY-QHRL-VSQFRHDPYKFLMRNHRKPKDYSNDMLLHLSEPADITDGVKVIDLP 120
 Qy 90 NKDGPEQEIPIVQSIHPHCYNS---SD-V---EDNHDMLLQLRDQASLGSKVKPISLA 142
 Db 121 TKEPKVSTCLVSWGS-TNPSNEFFDDLOCVNIHLLSNEKIKAYKEKVTDLMLCAGE 179
 Qy 143 DHCTQPGQKCTVSGWGTVTSRE-NFPDTLNCAEVKIPFQKKCEDAYPGQITDGMVCAGS 201
 Db 180 LEGKDTCRGDSGGPLICDGLGITSWGSVPCGPNKPGIYTKLIKFTSWIKVEM 235
 Qy 202 SKGA-DTCQDGGGFLVCDGALGITSWGSDFGSRDKPGVITNICRYLDWIKKII 256

RESULT 15

ID Q92046 PRELIMINARY; PRT; 249 AA.
 AC Q92046; O13261;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DE PREPROTEIN PRECURSOR (EC 3.4.21.4) (ALPHA- AND BETA-TRYPSIN).
 OS DISSOSTICHUS MAWSONI.
 OC EUKARYOTA; ANIMALIA; METAZOA; CHORDATA; VERTEBRATA; OSTEICHTHYES;
 OC ACTINOPTERYGII; PERCIFORMES; NOTOTHENIOIDEI; NOTOTHENIIDAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-PANCREAS;
 RA CHEN L., DEVRIES A.L., CHENG C.H.C.;
 RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA CHEN L., DEVRIES A.L., CHENG C.H.C.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 94:3811-3816(1997).
 CC -I- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-|-XAA, LYS-|-XAA.
 DR EMBL; U58945; G1399811;
 DR EMBL; U58835; G2102686; -;
 SIGNAL; HYDROLASE.
 SIGNAL 15 POTENTIAL.
 FT CHAIN 16 249 PROTRYPSIN.
 SQ SEQUENCE 249 AA; 27169 MW; 55C59EB4 CRC32;

Query Match 34.1%; Score 663; DB 12; Length 249;
 Best Local Similarity 41.8%; Pred. No. 5.33e-154;
 Matches 104; Conservative 49; Mismatches 86; Indels 10; Gaps 8;

Db 6 LLLLGAAAVPR-EDGRIIGGYECSPHSRPYMASLNYGHF-CGGVLLNNOWVLSVAHC 63
 Qy 15 LLLLGAAWAGSHRAQEKVLGHECQPHSQWQAALFQGGVLLCGVLYGGNWLTAACH 74
 Db 64 WNPYSQVILGDHNLRFECTEQMLKNTIIMHPSYDIQTL-DF--DIMLIKLYHFVEV 120
 Qy 75 KPKYKT--VRLGDHSLQKNDGPEQEIPIVQSIHPHCYNSDVEDNHDMLQLRDQASL 132
 Db 121 TEAVAPLPTSCPYGLSCVSGWGIKLGGEAYMTLLOCLNVPVDOQVCEYTPGL 180
 Qy 133 GSKYKPISLADHCTQPGQKCTVSGWGTVTSRENF-PDTLNCAEVKIPFQKKCEDAYPGQ 191
 Db 181 ISTTWACAGYMEGGKACNGDSGPLVCDGEGVGLVSWG-QGCAEPNPGVYVVKLCFFHS 239
 Qy 192 ITDGMVCAGSKGA-DTCQDGGGFLVCDGALGITSWGSDFGSRDKPGVITNICRYLD 250

Db 240 WFEVLAAN 248
 Qy 251 WIKKIIGSK 259

Search completed: Wed Jul 22 15:31:10 1998
 Job time : 76 secs.